AT-RISK GRADE 1–3 STUDENTS' UNDERSTANDING OF THE NUMBER SEQUENCE AND THE NUMBER LINE

Guri A. Nortvedt

University of Oslo

This paper presents the outcomes of a research project aimed at developing mapping tests designed to identify students at-risk of lagging behind in mathematics. Data from the first test implementation (N1 = 2,370, N2 = 2,483 and N3 = 2,286) are used to investigate at-risk grade 1–3 students' understanding of the number sequence and the number line, as well as to discuss how this knowledge develops across grades. Analyses indicate at-risk students demonstrate a weak understanding of the number sequence. When attempting to identify target numbers on a structured number line, many consistently count by one from zero. Evidence suggests at-risk students do not master counting by two, five or ten. In addition, some struggle with the conventions for writing numbers. More growth is seen from grade 1 to 2 than from grade 2 to 3.

INTRODUCTION

Norwegian grade 4 students scored at the international average in the 2011 Trends in Mathematics and Science Study (TIMSS 2011) (Mullis, Martin, Foy, & Arora, 2012). However, in total, 9% of students did not reach the low benchmark and only demonstrated some basic mathematical knowledge, such as adding and subtracting whole numbers. Still, the number of students reaching the low benchmark has risen substantially from 75% to 91%, from 2003 to 2011 (Mullis et al., 2012, p. 93). Previous studies have demonstrated Norwegian primary school teachers have had a tendency to 'wait and see' when they observe students with difficulties (Nordahl & Hausstätter, 2009), as well as that the majority of special needs education is given in lower secondary school (Solli, 2005). Consequently, intervention has often been delayed until students have developed real difficulties or lagged substantially behind. This pattern has worried the Norwegian Ministry of Education and Research, and in 2006, they released the white paper 'Early Intervention for Lifelong Learning', presenting a national policy for "how the education system can make a greater contribution to social equalisation" (MER, 2006, p. 1). A key aspect of the early intervention policy is to ensure every student acquires the basic skills needed to succeed in the educational system, and in 2008, a mandatory national screening test for grade 2 was introduced, followed by optional grade 3 and 1 testing in 2009 and 2011. The purpose of the implementation was to aid teachers in identifying students who had not acquired the basic mathematical concepts and computational skills necessary to provide a solid foundation for further learning. Based on the TIMSS 2011 outcomes, the strategy might be seen as successful. However, the majority of special needs education continues to be given at the lower secondary level (NDET, 2014), indicating teachers still either tend to wait and see or struggle to plan and carry out successful interventions for identified students.

In a What Works-report, Gersten et al. (2009) recommend screening all students to identify those at risk of potential difficulties in mathematics and to provide intervention to identified students. However, for the assessment to be a starting point for mathematical learning, teachers need to be able to use the test results to identify where learners are and where to go next (Wiliam, 2007). Consequently, assessments should be targeted to the student group of interest, and when the second generation of the mapping tests were developed, this was taken more into consideration. The assessment launched in April 2014 is a hybrid between a screening and a mapping test. While the full student cohort takes the test, it is targeted towards the weakest 20% of students (labelled 'at-risk students') and is developed to differentiate more securely between the first and second quintile groups of students (NDET, 2011). Consequently, the tests have a ceiling effect by design and consist of many easy items. A student close to the cut-off score typically solves 70–85% of the items correctly, which provides teachers with much more information about what students can do than assessments targeted at the full student population, where struggling students are characterised by what they cannot do.

This paper draws on experiences and data from the first test implementation in the spring of 2014. The research questions for the paper include: 1) what do the mapping tests display about grade 1–3 students' understanding of the number sequence and the number line and 2) how does this knowledge develop from grade 1 to grade 3.

UNDERSTANDING THE NUMBER LINE

Young children's knowledge of counting and quantity can be seen as a starting point for their understanding of numbers (Griffin, 2003). As they become more experienced with counting and numbers, their conceptual structure for whole numbers might be seen as a mental number line (Griffin, 2003). This structure functions as a mental counting line and allows students to compare the magnitude of numbers and understand place value. In school, students encounter the number line as a mathematical object representing a sequence of numbers, often starting from 0, representing only the positive numbers. However, a number line might start from a different number and have different spacing. Students need to understand how intervals with the same magnitude represent the same numerical difference. While the number line can be seen as a measurement object or a graphical representation of number, students often conceive it as a counting object (Diezmann & Lowrie, 2006). The empty number line might be used as a learning tool towards developing strategies for addition and subtraction and is often advocated in teacher education literature (for a Norwegian example, see Heiberg Solem, Alseth, & Nordberg, 2009). However, teachers might make ambiguous connections between the number line and the number track (Grey & Doritou, 2008). As a consequence, many students focus on the number line as a counting tool for handling addition and subtraction.

Poor understanding of the number sequence might hinder students' mathematical development, as this most likely signals a weak conceptual understanding of numbers. When assessing students' number concepts, number line items are frequently used as a representation of the number sequence (Diezmann & Lowrie, 2007).

TEST DEVELOPMENT AND TECHNICAL REQUIREMENTS

The Norwegian national mapping tests implemented in 2014 are designed to measure students' number concepts and calculation skills. According to the test development framework provided by the Norwegian Directorate for Education and Training (NDET, 2011), tests should serve as a formative assessment tool for the first quintile group (the 20% weakest students), in that test results should function as a starting point for further classroom work. The class teacher administers tests towards the end of grades 1–3 when students are six, seven and eight years old. While the grade 2 test is mandatory, tests for grades 1 and 3 are optional. The school principal or the local school administration decides on participation. Tests are scored by the teacher.

One test was developed for each grade level drawing on research literature (for instance, Geary, 2004; Griffin, 2003; Verschaffel, Greer, & De Corte, 2007), and pre-tested in cognitive labs and through small group testing before being pilot tested in spring 2013. Test reliability, measured by Cronbach's alpha coefficient, is high and larger than .93 for all tests. In addition, no gender DIF was observed. As described in the test framework (NDET, 2011), all tests discriminate around a 20% cut-off score, which is between the first (Q1) and second (Q2) quintile groups. This means that for most test items, 70–90% of students successfully solve it. All tests have a ceiling effect by design. Tests are timed to identify students with naive or rigid counting strategies.

All tests were validated using different expert panels (primary school students, researchers, test designers, teachers, special education teachers, school leaders and international experts). Experts commented on single items, clusters of items and full tests. All experts agreed that the items described below assess some aspect of understanding of the number sequence.

Table 1. Test characteristics, means and number of students in Q1 and Q2.

Grade	No. items	Cut-off	Mean Q1	n Q1	Mean Q2	n Q2
1	50	39	36.3 (7.697)	454	42.4 (1.401)	448
2	55	41	31.7 (8.720)	500	45.6 (1.889)	541
3	72	59	48.7 (10.089)	443	63 (1.704)	488

Sample

The dataset used from this paper was collected during the first test implementation. A representative national sample was drawn, excluding schools with fewer than five students at the grade level and international schools. In total, 127 schools participated, with one student group/class at each grade level: N1 = 2,370, N2 = 2,483 and N3 = 2,286. Only data from Q1 and Q2 students are used for this paper (see Table 1).

Items assessing students' understanding of the number line

Items assessing students' understanding of the number line were developed for each grade level, albeit with a different range of numbers: 0–20 for grade 1, 0–100 for grade 2 and 0–300 for grade 3. Four types of items will be discussed in this paper (the number of items for each grade level used in this paper is given in parenthesis).

- Counting objects and indicating on a labelled number line how many (3/2/2)
- Placing numbers on a number line (2/0/2)
- Counting on from a given number (counting up and down) (4/5/4)
- Sorting numbers by magnitude (3/3/3)

Examples will not be given, nor will the numbers used in the items, although their magnitude will be indicated. As the test items are not released, items will only be described in the results section. Response rates will be reported for single items and groups of items for the two lowest quintile groups. Due to the test design, significant differences are observed between Q1 and Q2 in all grades for all four groups of items.

RESULTS

Tests designed to identify the weakest 20% of students could have been aimed at the average achievers, applying items that the weak students do not master at all. However, teachers must be able to draw conclusions about the competences of their students from student response patterns and identify weaknesses and strengths in their students' mathematical competences for the test to function as an assessment for learning for the targeted student group (Wiliam, 2007). Consequently, the tests mainly comprise items that are very easy to most students. In addition, Q2 students have a generally higher probability then Q1 to solve correctly all items. Significant differences were observed between Q1 and Q2 for all groups of items for all grade levels as measured with ANOVA, for grade 2 for instance $F(2,2479) = 386.562$, $p < .01$, $F(2,2480) = 1114.889$, $p < .01$ and $F(2,2480) = 528.767$, $p < .01$ for items group 1, 3, and 4 respectively.

The first group of number line items asked students to count concrete objects and mark the corresponding number on the number line by drawing a line from the group of objects to the location of the number. Figure 1 displays a marked number line similar to the number lines used for grade 1 test items. Students were asked to count small, unstructured groups of objects (maximum 15), and 56–82% of Q1 students successfully solved these items. Items discriminated satisfactorily between the two first quintile groups when students were asked to count more than 10 objects. A difference in percentage points ranging from 9 to 27 between the two groups indicated Q2 mastered the items to a much larger extent than Q1.

Q1 students took longer to finish these items. In total, 16% of students in this group did not manage to finish the last item, compared with 2% of the Q2 students. Pencil marks in the test booklets indicated many of the struggling students counted by 1 from 0 to identify the numbers on the number line. Figure 1 displays how one student counted to identify 6, 9 and 12 on a similar item, used in test development. Such marks were frequently observed in test booklets. It is hypothesised that the weakest students solve such items by double counting; first the objects followed by counting up to the target number on the number line.

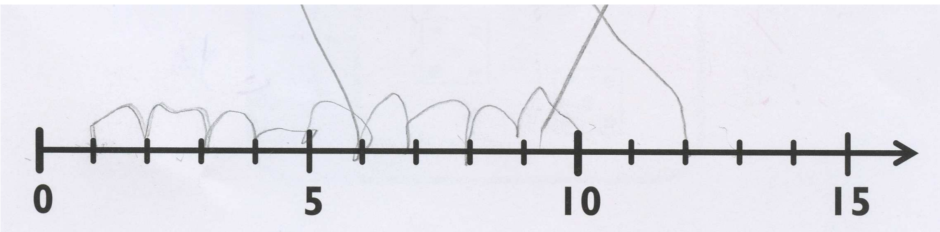

Figure 1. Student work on a structured number line, grade 1.

To monitor growth across grade levels, a set of identical items was given to students in grades 2 and 3. Students were asked to count structured groups of objects (grouped in tens and single units). Items included up to 39 objects (one, two or three groups of 10 and some single objects). Number lines were structured, and each whole number was marked. In addition, the structure of the number sequence was given either by labelling every 5 and 10 or by labelling every 10 only. Approximately half the Q1 students solved the items correctly in both grades, indicating these tasks are still challenging to the weakest students, even after three years of being exposed to these types of items from their textbooks. Again, students counting by 1 from 0 were observed by pencil marks in the test booklets. In addition, fewer Q2 students mastered the items in grade 3 compared to grade 2. However, grade 3 students were allowed a shorter time to solve this item, which might explain this observation.

The second group of items asked grade 1 and 3 students to mark on the number line the position of some given numbers. Behaviours similar to those used for the previous group of items were observed for some Q1 students who apparently counted by 1 and ignored the marking and sequencing of the number line. This strengthens the hypothesis that many Q1 students do not understand the number line structure. As the magnitude of the numbers, the marking and the labelling of the number lines were different for the number line items for grades 1 and 3, direct comparisons of difficulty level cannot be made. It should be noted that only number lines that marked every whole number were used for grade 3, thus enabling students who count securely to identify the correct position of a number, provided they started counting from the correct number and position. When assigning the starting point of the grade 3 number lines at a number other than 0, some Q1 students still started counting from 1 as if all number lines start at 0. For struggling grade 1 students, numbers larger than 10 where challenging and a marked difference was between Q1 and Q2, with 40% and 74% of students, respectively, solving these items correctly. Development is seen across grade level, and Q1 grade 3 students confidently handled numbers smaller than 50. However, numbers larger than 100 were still challenging.

The third group of items reported in this paper were items that required students to recognise number patterns and count on from a given number. The mapping tests are paper-and-pencil tests, and students had to provide written responses to the count-on items. Consequently, these items also provided information about students' knowledge of the conventions for writing numbers (e.g. that fourteen is written 14 and not 41). Students were typically given three numbers and asked to count on from the last. Numbers ranged from 0–15 (grade 1) and 0–70 (grades 2 and 3). Grade 1 at-risk students confidently counted up but struggled to count down. Counting down was challenging, even to Q2 students, with approximately 42% and 70%, respectively, managing to count down from a number smaller than 15. Grade 1 students were mainly asked to count by 1. Counting by two differentiated mainly between the second and third quintile groups; only 14% and 34% of Q1 and Q2 students, respectively, solved this item correctly in grade 1. Counting by two, five and 10 can be efficient for executing mental calculations quickly, but this rests on secure knowledge of the number sequence. This is one of the areas where the Q1 and Q2 differ substantially in grades 2 and 3. While 80% and 93% of Q2 students on average solve such items, respectively, 45% and 59% of Q1 students on average answered the same items correctly, respectively, indicating weaker knowledge of the number sequence. Q1 students also took longer to solve these items, and in grade 2, as many as 30% did not manage to finish the last item within the given time, supporting the hypothesis of a weaker understanding of the number line.

The last group of items required students to sort a given set of numbers by magnitude (up to 5 numbers). To sort and compare successfully, students need a well-developed mental number line (Griffin, 2003). Grade 1 students were asked to sort numbers in the range of 0–20, grade 2 in the range of 0–100 and grade 3 in the range of 0–300. While more than 80% of Q2 students at all grade levels successfully sorted the numbers, this is one area where much development in seen from grade 1 to grade 3 among Q1 students. In grade 1, numbers larger than 10 represent a challenge and only 55% of students in Q1 managed to sort five numbers, of which four were larger than 10, indicating almost half the students identified as being at-risk struggled to understand the magnitude and ordering of small, positive whole two-digit numbers. Q1 students in grade 2 struggled with items where they needed to compare numbers, such as ab and ba or ab, ca and ad. A few students were likely unable to handle comparing as many numbers as five, and responses indicated these students probably compared pair-wise or a few numbers at a time when sorting. In addition, at-risk students took longer to complete these items.

The numbers given to the grade 3 students masked some of the difficulties displayed by grade 2 students in relation to not understanding place value or not mastering the conventions of writing numbers. However, students were also observed to struggle when comparing numbers like abc and aca. Still, many at-risk students struggled with sorting items, and of the six items given to grades 2 and 3, between 41% and 86% of the students, respectively, sorted the numbers correctly.

DISCUSSION AND CONCLUDING REMARKS

In danger of a circular argument due to the test design, analyses of student response patterns on the mapping tests revealed that Norwegian at-risk primary school students displayed difficulties and a lack of conceptual understanding and procedural use of the number line, as anticipated from previous research. Q1 students also demonstrated a weak knowledge of the number sequence, as might be anticipated from the number line items. A further analysis demonstrated Q1 students took longer than Q2 students to count up the number line, especially when only end points were labelled. Items demanding 'double' counting, as in the first group of items, which asked students to count small numbers of concrete objects and tie this amount to the number line, were more challenging to Q1 than to other students. It might be assumed more at-risk students started counting from 0 on the number line when solving these items rather than using the numbers identified by the labelling of the number line to navigate. It could be argued that these students view the number line as a counting tool rather than a measurement tool, as discussed by Diezmann and Lowrie (2006, 2007). However, their knowledge of the number sequence, displayed by the counting-on and sorting of items, is also weak, and it might also be assumed they have less support from the structuring of the number lines. For instance, some weak students struggled with the conventions of written numbers, indicating a lack of understanding of place value.

Learning the number sequence and understanding place value and the number line are crucial to learning basic arithmetic and are essential to the primary school curriculum. Consequently, items assessing these basic concepts must be included in a mapping test if the teacher is to use the test for formative purposes. Traditionally, at-risk students have been identified by their strategy use when solving simple addition and subtraction tasks (see for instance Geary, 2004; Griffin, 2003). However, number line items are well suited to identify these students and to provide valuable information on students' number concepts that can serve as a starting point for teaching interventions.

Analyses of student responses to the mapping tests revealed Norwegian teachers might need further support to target teaching to their students' needs, as the same use of simple counting strategies was found among students in all three grade levels; this indicates many weak students do not overcome their difficulties through existing teaching activities. In addition, error patterns might indicate weak students have a much slower development of a conceptual understanding from grade 1 to grade 3 than typical developing students, as might be expected (see, for instance, Geary, 2004). While a fairly strong development was observed from grade 1 to grade 2, little improvement was seen from grade 2 to grade 3. The weak grade 1 students most likely do not catch up with other students during grade 2.

References

Diezmann, C. M., & Lowrie, T. (2006). Primary students' knowledge of and errors on number lines. In P. Grootenboer, R. Zevenbergen & M. Chinnappan (Eds.), *Proceedings of* *the 29th annual conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 171-178). Sydney: MERGA.

Diezmann, C. M., & Lowrie, T. (2007). The development of primary students' knowledge of the structured number line. In J. H. Woo, H. C. Lew, K. S. Park, & D. Y. Seo (Eds.), *Proceedings of the 31st conference of the International Group for the Psychology of Mathematics Education*. (Vol. 2, pp. 201-209). Seoul: PME.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37*(1), 4-15.

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Responses to intervention (Rtl) for elementary and middle schools* (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. http://ies.ed.gov/ncee/wwc/publciations/practiceguides/

Grey, E., & Doritou, M. (2008). The number line: Ambiguity and interpretation. In O. Figueras, J. L. Cortina, S. Alatorre, T. Rojano & A. Sepúlveda (Eds.), *Proceedings of the joint meeting of PME 32 and PME-NA XXX* (Vol. 3, pp. 97-104). Morelia: PME.

Griffin, S. (2003). The development of math competence in the preschool and early school years: Cognitive foundations and instructional strategies. In J. M. Royer (Ed.), *Mathematical cognition* (pp. 1-32). Greenwich, CT: Information Age.

Heiberg Solem, I., Alseth, B., & Nordberg, G. (2009). *Tall og tanke. Matematikkundervisning på 1. - 4. trinn [Numbers and thoughts. Mathematics teaching in grades 1 - 4]*. Oslo: Gyldendal Akademisk.

Ministry of Education and Research (MER). (2006). *White paper 16, 2006–2007*.

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International Results in Mathematics. Boston and Amsterdam:* TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Norwegian Directorate for Education and Training (NDET). (2014). *The education mirror 2014.* Oslo: NDET.

NDET. (2011). *Rammeverk for kartleggingsprøver på barnetrinnet [Framework for primary school mapping tests]*. Oslo: NDET.

Nordahl, T., & Hausstätter, R. S. (2009). *Spesialundervisningens forutsetninger, innsats og resultater [The conditions, efforts and results of special needs education]*. Hamar: Hamar University College.

Solli, K. A. (2005). *Kunnskapsstatus om spesialundervisning i Norge [Knowledge about special needs education in Norway]*. Oslo: Directorate for Education and Training.

Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. K. J. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 557-628). Charlotte, NC: Information Age.

Wiliam, D. (2007). Keeping learning on track. In F. K. J. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 1053-1098). Charlotte, NC: Information Age.